Filozofia w Praktyce
Tom 2 (2016), Artykuł 7

AUTONOMICZNE SYSTEMY BOJOWE A KONFLIKTY ASYMETRYCZNE

– Michał Klincewicz –

Autonomiczne systemy bojowe (w skrócie „AWS", od ang. *Autonomous Weapon Systems*) same wybierają cele i są w stanie autonomicznie, to znaczy bez ingerencji człowieka, otworzyć ogień. Czy wprowadzanie do sił zbrojnych coraz bardziej zaawansowanych AWS (takich jak Samsung SGR-A1 i ATK PAWS lub nowy rosyjski czołg Uran-9) powinno wzbudzać nasz entuzjazm? Ronald Arkin, pionier w badaniu robotów bojowych, twierdzi że tak (Arkin 2010). Przekonuje, że AWS są odporne na stres, zmęczenie lub ból, dzięki czemu nie zabiją umyślnie cywilów, nie złamią rozkazów i nie wrócą z wojny z zespołem stresu pourazowego. Jeżeli tak, to wdrożenie AWS byłoby nie tylko moralnie dopuszczalne, ale też nakazane. Ich wprowadzenie powinno wpłynąć na zmniejszenie potwornych, z punktu widzenia moralności, skutków wojny.

Niedawno opublikowałem artykuł, w którym na pytanie o moralną dopuszczalność wprowadzenia AWS odpowiadam negatywnie (Klincewicz 2015). Moim zdaniem Arkin ma rację, kiedy twierdzi, że AWS mogą z powodzeniem w przyszłości zastąpić ludzi na polu walki oraz okazać się lepszymi od nich żołnierzami. Jednakże wprowadzenie tego rodzaju rozwiązań związane jest z ryzykiem, które stawia pod znakiem zapytania ich moralną dopuszczalność. Upraszczając, mój argument wygląda następująco:

1. Skomplikowane oprogramowanie nieuchronnie zawiera błędy (*bugs*).
2. Błędy stanowią słabe punkty (*vulnerabilities*) oprogramowania.
3. Słabe punkty ułatwiają zhakowanie oprogramowania.
4. AWS mają (i będą mieć) ogromnie skomplikowane oprogramowanie.

Zatem:

5. AWS mogą zostać zhakowane.
6. Zhakowane AWS doprowadzą do konsekwencji znacznie bardziej moralnie odrażających niż te, które mogą być spowodowane przez żołnierzy.

Zatem:

7. Używanie AWS w wojnie jest moralnie niedopuszczalne.

Z każdą przesłanką powyższego argumentu można polemizować. Z jednej strony, przesłanki (1) – (5) są uzależnione od falsyfikowalnych korelacji, które mogą się zmienić z czasem lub wynikać z błędu statystycznego. Z drugiej – są one jednak w pewnym stopniu uzasadnione i wykorzystują argumentację, z którą można się spotkać w informatyce. Przesłanka (6) jest natomiast czystą spekulacją i wymaga obrony. Nie przeprowadziłem takiej obrony w swoim artykule, skorzystam zatem z okazji, żeby przeprowadzić ją tutaj.

Pierwszy problem z (6) to założenie, że o moralnej dopuszczalności lub obligatoryjności decydują konsekwencje wdrożenia AWS. Po drugie, nawet jeżeli założymy że

tak jest, to nie jest jednak oczywiste, dlaczego powinniśmy przyjąć, że zhakowany AWS może spowodować więcej szkody niż żołnierze. Pierwszy z tych problemów nazwijmy „problemem teoretycznym", a drugi – „problemem empirycznym".

Problem teoretyczny jest dość łatwy do przezwyciężenia. Przypomnijmy sobie argument Arkina, który twierdzi, że powodem, dla którego jest dopuszczalne lub nawet nakazane, by wdrożyć AWS jest eliminacja złych konsekwencji powiązanych z zachowaniem żołnierzy w czasie wojny i po niej. W odróżnieniu od żołnierzy, AWS nie będą miały urazów, nie będą podatne na stres oraz nie będą ignorować rozkazów, a zatem ich wdrożenie wyeliminuje źródła negatywnych *konsekwencji*. Tymi konsekwencjami mogą być naruszenia praw wojny, w tym naruszenia konwencji genewskich, lub zasad regulujących użycie siły i broni przez żołnierzy podczas misji. Dla jakości tej polemiki jest ważne, by mój argument akceptował przesłanki argumentu Arkina, więc także i jego rachunek konsekwencji. Pozwala to bowiem na przeprowadzenie rzetelnej krytyki.

Podsumowując, problem teoretyczny nie stanowi trudności dla mojego argumentu, ponieważ odwołanie się do konsekwencji oznacza po prostu akceptację argumentu Arkina w jego najmocniejszej formie. Jeśli ktoś odrzuca rachunek konsekwencji jako probierz dopuszczalności moralnej, może na tej samej podstawie za jednym zamachem odrzucić również argument za dopuszczalnością wdrożenia AWS.

Problem empiryczny jest jednak trudniejszy do rozwiązania. Nie jest bowiem jasne, w jaki sposób moglibyśmy argumentować z jakimkolwiek stopniem pewności, że zhakowane AWS doprowadzą do jeszcze bardziej moralnie odrażających konsekwencji od tych, które byłyby w identycznych warunkach spowodowane przez żołnierzy. Możliwym źródłem spekulacji na ten temat jest odwołanie się do wydarzeń z przeszłości. Jednak nie ma w aktualnej przeszłości przypadków wykorzystania AWS, na których można by opierać prognozy dotyczące przyszłości. W tej sytuacji, jedynym źródłem przewidywania mogą być analogie.

Możliwa analogia odwołuje się do pozyskania broni jądrowej, czego efektem była doktryna Wzajemnego Zagwarantowanego Zniszczenia (w skrócie MAD z angielskiego *Mutual Assurred Destruction*) (Morris, 2009). Konsekwencją MAD był okres względnego pokoju, w trakcie którego światowe potęgi nie prowadziły ze sobą otwartych wojen w obawie przed zagładą nuklearną. W momencie kiedy mocarstwa militarne zostałyby wyposażone w AWS, rzeczywiście zaistniałaby możliwość długotrwałych rezultatów podobnych do efektu MAD, tzn. pokoju i stabilności. To jednak przeczy przesłance (6). Przypominam, że głosi ona, że zhakowane AWS doprowadzą do konsekwencji znacznie bardziej moralnie odrażających niż te, które mogą być spowodowane przez żołnierzy.

Istnieją jednak dalsze analogie. Jedną z konsekwencji MAD było to, że mocarstwa zaczęły angażować się w konflikty z przeciwnikami o całkowicie nieproporcjonalnych możliwościach. Innymi słowy, wojny stały się asymetryczne. Najbardziej istotną cechą asymetrycznego konfliktu jest sposób prowadzenia walki przez słabszą stronę:

> [Słabsza strona] przeprowadza ataki na niższym poziomie taktycznym w nadziei, że wywrą one duży wpływ na wyższym poziomie strategicznym, są to na przykład: strategicznie detonowane bomby powodujące zmianę polityki rządu strony dominującej, atak hakerski na jej system komputerowy, który ma negatywny efekt na całą

> gospodarkę, strącenie jednego samolotu, aby powstrzymać całe naloty bombowe, usunięcie jednego okrętu, które ma zatrzymać całą armadę, zabicie kilku żołnierzy, co powoduje »odwrót« oraz przeciągnięcie się wojny w czasie po to, żeby społeczeństwo straciło nią zainteresowania czy chęć wygrania (Thornton 2007, s. 22).

W tym miejscu analogia z konsekwencjami zakładanymi przez (6) jest wyraźna. Wszystkie AWS ta trakcie asymetrycznego konfliktu stanowią doskonały cel dla ataku hakerskiego na niższym poziomie taktycznym, który da potencjalnie ogromny wpływ na wyższym poziomie strategicznym. Co się stanie, jeżeli zhakowane AWS zaczną strzelać do ludności cywilnej? Co się stanie, jeżeli zhakowane AWS zaczną atakować własne bazy? Z całą pewnością odniesie to spory efekt na poziomie strategicznym. Skuteczność tego rodzaju akcji zwielokrotni możliwości stosowania strategii opisanych przez Thorntona w zacytowanym fragmencie.

Słabsze strony w asymetrycznych konfliktach, które tak naprawdę będą jedynymi przeciwnikami AWS, dołożą wszelkich starań do tego, aby wykorzystać luki w oprogramowaniu AWS do maksymalizowania efektywności na poziomie strategicznym. Mogą wykorzystać choćby medialny potencjał przypadków zhakowania AWS oraz ich potencjalnie śmiercionośne efekty. Dlatego należy się spodziewać, że wdrożenie i ewentualne zhakowanie AWS doprowadzi do konsekwencji znacznie bardziej moralnie odrażających niż te, które mogą być spowodowane przez żołnierzy. AWS nadają się do tego idealnie.

Literatura

Arkin, R.C. (2010), *The Case for Ethical Autonomy in Unmanned Systems*, „Journal of Military Ethics" 9 (4): 332–341. doi:10.1080/15027570.2010.536402

Klincewicz, M. (2015), *Autonomous Weapon Systems, the Frame Problem, and Computer Security*, „Journal of Military Ethics" 14 (2): 162–176. doi:10.1080/15027570.2015.1069013

Morris, C. (2009), *Kontraktualna obrona odstraszania nuklearnego*, [w:] *Etyka wojny. Antologia*, T. Żuradzki, T. Kuniński (red.), Wydawnictwo Naukowe PWN, Warszawa.

Thornton, R. (2007), *Asymmetric Warfare: Threat and Response in the 21st Century*. Polity Press, Oxford.

Michał Klincewicz – adiunkt w Zakładzie Kognitywistyki Instytutu Filozofii UJ.

Tekst powstał dzięki finansowaniu z Fundacji na rzecz Nauki Polskiej na podstawie Umowy nr 139/UD/SKILLS/2015 o wykorzystanie Nagrody przyznanej w konkursie eNgage w ramach projektu SKILLS współfinansowanego z Europejskiego Funduszu Społecznego.